

On Optimal Input Design for Feed-forward Control

Per Hägg and Bo Wahlberg

Abstract—This paper considers optimal input design when the intended use of the identified model is to construct a feed-forward controller based on measurable disturbances. The objective is to find a minimum power excitation signal to be used in system identification experiment, such that the corresponding model-based feed-forward controller guarantees, with a given probability, that the variance of the output signal is within given specifications. To start with, some low order model problems are analytically solved and fundamental properties of the optimal input signal solution are presented. The optimal input signal contains feed-forward control and depends of the noise model and transfer function of the system in a specific way. Next, we show how to apply the partial correlation approach to closed loop optimal experiment design to the general feed-forward problem. A framework for optimal input signal design for feed-forward control is presented and numerically evaluated on a temperature control problem.

I. INTRODUCTION

SYSTEM identification is about the estimation and validation of mathematical models of dynamical systems from experimental data. It is well recognized that the input signal used to excite the system during the experiment significantly affects the accuracy of the identified model. A correctly chosen excitation signal can immensely improve the quality of the resulting model while a poorly chosen signal can result in a useless model. This motivates input or experiment design in system identification.

The quality of a model depends on the intended use of the model. A good experiment should highlight the important properties for the intended application. This is the main idea of *identification for control* [1], [2], *least costly identification* [3] and *applications oriented input design* [4].

Much work has been focused on input design when the model is to be used in a control application, see for example [5] or [6] for application to MPC. However, many industrial control systems do not only utilize feedback but also feed-forward control. The idea is to measure a disturbance, predict its impact on the plant and then compensate for it with the input signal. Compensating for measurable disturbances by feed-forward control can improve the performance considerably compared to when only feedback control is used. Measurable disturbances could for example be the outside temperature when controlling the indoor temperature in a house.

This work was partially supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC Grant Agreement No. 267381, the Swedish Research Council and the Linnaeus Center ACCESS at KTH

The authors are with the Automatic Control Lab and ACCESS, School of Electrical Engineering, KTH, The Royal Institute of Technology, SE-100 44 Stockholm, Sweden.

E-mail: {pehagg, bo}@kth.se

The aim of this paper is to study optimal input design when the intended use of the model is for feed-forward control. We will use the application oriented input design framework presented in [4]. To illustrate how system properties affect the optimal input signal we start by analyzing a problem which can be solved analytically. We then present a framework for input design for more general systems and show how to formulate them as Semi Definite Programs (SDP) that can be solved efficiently using numerical methods.

The outline of the paper is as follows. In Section II we define the input design problem and give some preliminary results from application oriented input design. Section III analyzes first order FIR subsystems and make some observations. A more general framework for input design for feed-forward control is given in Section IV. The framework is then applied to a simulation example in Section V while Section VI concludes the paper.

II. PROBLEM FORMULATION

Consider the system in Figure 1. The dynamics of the

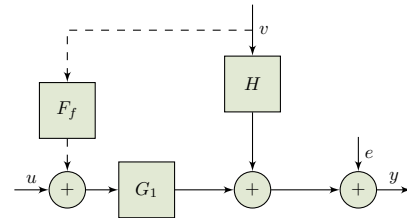


Fig. 1: The considered system. The identification setup without feed-forward filter (F_f) and application setup with feed-forward filter.

system can be written as

$$y_t = G_0(q)u_t + H_0(q)v_t + e_t,$$

where q^{-1} is the delay operator, i.e., $q^{-1}u_t = u_{t-1}$. Here $G_0(q)$ and $H_0(q)$ are the plant transfer functions. The input to the system is u_t while y_t is the measured output disturbed by a zero mean white noise signal e_t with variance λ_e . The process v_t corresponds to a measurable disturbance to the system. We assume that v_t is a stationary stochastic process with known spectral properties, i.e., v_t can be written as $v_t = M(q)s_t$ where s_t is a zero mean Gaussian process with variance λ_s and the stable and inversely stable filter $M(q)$ is known. The spectrum of v_t equals $\Phi_v(\omega) = |M(e^{j\omega})|^2 \lambda_s$.

The objective is to design a feed-forward controller F_f in Figure 1 to suppress the effect of the measurable disturbance in y_t . However, the dynamics of the subsystems H and G are assumed unknown and need to be identified. Hence we

want to identify the system dynamics and use the identified model to design a feed-forward controller to counteract the influence of the measured disturbance v_t on the output y_t . In particular, we will study to design the input u_t during the experiment such that we with high probability can guarantee that the designed feed-forward controller satisfy an accuracy constraint.

We parameterize the submodels as

$$G(q, \theta), H(q, \theta),$$

where $\theta \in \mathbb{R}^n$ is the model parameter vector and we assume that there is a vector θ_0 such that

$$G(q, \theta_0) = G_0(q), H(q, \theta_0) = H_0(q),$$

that is, the true system can be described by the model.

We denote the model parameter vector estimated using a Prediction Error Method (PEM), [7], from N data points of the inputs and outputs, $\{u_t, v_t, y_t, t = 1, \dots, N\}$ as $\hat{\theta}_N$. Under some mild assumptions the asymptotic (in N) covariance matrix of the estimated parameters

$$P = \lim_{N \rightarrow \infty} NE \left\{ (\hat{\theta}_N - \theta_0)(\hat{\theta}_N - \theta_0)^T \right\},$$

can be written as

$$P^{-1} = \frac{1}{2\pi\lambda_e} \int_{-\pi}^{\pi} \mathcal{F}_0(e^{j\omega}) \Phi_{\chi_0}(\omega) \mathcal{F}_0^*(e^{j\omega}) d\omega, \quad (1)$$

where the joint spectrum

$$\Phi_{\chi_0}(\omega) = \begin{bmatrix} \Phi_u(\omega) & \Phi_{us}(\omega) \\ \Phi_{us}^*(\omega) & \lambda_s(\omega) \end{bmatrix},$$

is the spectrum of $\chi_0 = [u_t \quad s_t]^T$ and

$$\mathcal{F}_0(q) = \left[\frac{\partial G(q, \theta)}{\partial \theta} \quad M(q)^* \frac{\partial H(q, \theta)}{\partial \theta} \right] \Big|_{\theta=\theta_0}, \quad (2)$$

see [7] for details.

Remark 1: Since we can measure v_t , we can correlate the input u_t with s_t (since we know $M(q)$) and hence Φ_{us} could be non-zero.

Further we know that, asymptotically, the identified parameter vector $\hat{\theta}_N$ lies within the set

$$\mathcal{E}_{SI} = \left\{ \theta : \frac{1}{2} [\theta - \theta_0]^T P^{-1} [\theta - \theta_0] \leq \frac{\kappa}{N} \right\}$$

with probability α . The constant κ can be determined from the $\chi^2(n)$ distribution as a function of the probability α and n , the number of estimated parameters.

A. Application oriented input design

As mentioned earlier we would like to identify the parameters of the system and from this design a feed-forward controller. The feed-forward controller will be denoted $F_f(q, \theta)$.

Using the framework in [4], we let $V_{app}(\theta)$ denote an application cost that measures the degradation in performance due to model errors. Here the application cost will represent the output variance from the measured disturbance v_t when designing the feed-forward controller with the identified

parameters instead of the true parameters, *i.e.*, all acceptable models satisfy

$$\begin{aligned} V_{app}(\theta) &= E \{ (H_0 + G_0 F_f(\theta)) v_t \}^2 \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_0 - G_0 F_f(\theta)|^2 \Phi_v d\omega \leq \frac{1}{\gamma}, \end{aligned}$$

where γ is some positive constant. The requirement is hence that the output variance should be less than $1/\gamma$. The overall objective of the input design is thus to design our input signal to be used during the experiment such that $V_{app}(\hat{\theta}_N) \leq 1/\gamma$.

We can approximate the application cost by a Taylor series expansion

$$V_{app}(\theta) \approx \frac{1}{2} [\theta - \theta_0]^T V''_{app} [\theta - \theta_0],$$

and hence the set of all parameters, θ that satisfies $V_{app}(\theta) \leq 1/\gamma$ can be approximated by the ellipsoidal set

$$\mathcal{E}_{app} = \left\{ \theta : \frac{1}{2} [\theta - \theta_0]^T V''_{app} [\theta - \theta_0] \leq \frac{1}{\gamma} \right\}. \quad (3)$$

B. Minimum Variance Input Design

To satisfy the application constraint we need to insure that the identified parameters lies within the application set (3) with high probability, *i.e.*, that $\mathcal{E}_{SI} \subseteq \mathcal{E}_{app}$. This is equivalent to

$$\frac{N}{\kappa} P^{-1} \succeq \gamma V''_{app}(\theta_0),$$

where $A \succeq B$ means that $A - B$ is positive semidefinite. If the above inequality holds then $\hat{\theta} \in \mathcal{E}_{app}$ with at least probability α .

The objective here is to find the minimum variance input signal used during the system identification experiment such that the identified model will satisfy the requirements from the application. More formally this can be stated as the following optimization problem

$$\begin{aligned} &\text{minimize}_{\Phi_u, \Phi_{ue}} \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) d\omega \\ &\text{subject to} \quad \frac{N}{\kappa} P^{-1} \succeq \gamma V''_{app}(\theta_0) \\ &\quad \Phi_{\chi_0} \text{ defines a spectrum.} \end{aligned} \quad (4)$$

Assume that we can express the constraint that Φ_{χ_0} defines a spectrum as an LMI-constraint. Then since in many cases, as we will see in the following, P^{-1} is an affine function of the spectral density Φ_{χ_0} and since $\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) d\omega$ is linear in Φ_{χ_0} , the optimization problem (4) becomes a SDP and can be solved efficiently. We will come back to how to write the spectrum constraint as an LMI later.

III. FIRST ORDER FIR SUBSYSTEMS

To get some insight in the problem we start by looking at a low order example where we can solve the optimization problem analytically. We assume that the two subsystems are first order FIR-systems, *i.e.*,

$$G_0(q) = 1 + b_0 q^{-1}, H_0(q) = 1 + h_0 q^{-1},$$

and the measurable disturbance is zero mean white noise with variance λ_v , *i.e.*, that $M = 1$ and $\lambda_v = \lambda_s$. We will use

system identification to identify the parameters $\theta = [b \ h]^T$ and use this to design a feed-forward controller.

The feed-forward controller will be on the form

$$F_f(q, \theta) = \frac{c_0 + c_1 q^{-1}}{1 + a_0 q^{-1}}, \quad (5)$$

where the constants c_0 , c_1 and a_0 are functions of h and b . Assuming that we know the true underlying system we design the optimal feed-forward controller, on the form (5), that minimizes the output variance due to the disturbance v_t . If $G(q)$ is minimum phase then obvious we should invert G in the feed-forward filter. In the non-minimum phase case it is known from minimum variance control that one should mirror the zero and then invert the system, see [8]. The optimal parametrization is thus given by

$$[c_0 \ c_1 \ a_0] = \begin{cases} [1 \ h \ b] & \text{if } |b| < 1 \\ \left[\frac{hb^2+b-h}{b^3} \ \frac{h}{b^2} \ \frac{1}{b} \right] & \text{if } |b| > 1. \end{cases}$$

The objective of the system identification is to find an estimate $\hat{\theta}_N$ of θ such that when they are used to design a feed-forward controller $F_f(\hat{\theta}_N)$ the output variance due to the disturbance is less than $1/\gamma$. All parameters, θ , that satisfies this is given by

$$V_{app}(\theta) = E \{ ((H(\theta_0) + G(\theta_0)F_f(\theta))v_t)^2 \} \leq \frac{1}{\gamma}. \quad (6)$$

A. Minimum phase system

First we look at the case when G_0 is minimum phase, i.e., $|b| < 1$. The application cost (6) can now be calculated, using for example residue calculus, to

$$V_{app}(\theta) = \frac{c^2 + d^2 - 2cdb}{1-b} \lambda_v \leq \frac{1}{\gamma},$$

where $c = h_0 - h + b - b_0$, $d = h_0 b - b_0 h$. The Hessian is then given by

$$V''_{app}(\theta_0) = 2\lambda_v \begin{bmatrix} p & -1 \\ -1 & 1 \end{bmatrix}, \quad p = \frac{h_0^2 - 2h_0 b_0 + 1}{1 - b_0^2}.$$

Using (1) the asymptotic covariance matrix for the identified parameters can be calculated as

$$P^{-1} = \frac{1}{2\pi} \int_{-\pi}^{\pi} I \begin{bmatrix} \Phi_u & \Phi_{uv} \\ \Phi_{uv}^* & \lambda_v \end{bmatrix} I d\omega = \begin{bmatrix} r_u(0) & r_{uv}(0) \\ r_{uv}(0) & \lambda_v \end{bmatrix},$$

where $r_u(k) = E[u_t u_{t-k}]$ and $r_{uv}(k) = E[u_t v_{t-k}]$ are the autocovariance of u_t and the cross covariance between u_t and v_t , respectively.

Since the input power can be written as

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) d\omega = r_u(0),$$

the optimization problem (4) can be reformulated as

$$\begin{aligned} & \underset{r_u, r_{uv}}{\text{minimize}} && r_u(0) \\ & \text{subject to} && \frac{N}{\lambda_e \kappa} \begin{bmatrix} r_u(0) & r_{uv}(0) \\ r_{uv}(0) & \lambda_v \end{bmatrix} \succeq \gamma \lambda_v \begin{bmatrix} p & -1 \\ -1 & 1 \end{bmatrix}. \end{aligned}$$

Note that condition that r_u and r_{uv} corresponds to a realizable experiment (or that Φ_{χ_0} defines a spectrum) is that the

matrix P^{-1} is positive semidefinite. However, since V''_{app} is positive semidefinite this is already taken care of in the optimization problem.

The optimization problem above can be solved analytically, see [9] for details. The optimization problem above is feasible if

$$\frac{N}{\lambda_e \kappa} - \gamma > 0$$

and the optimal solution is given by

$$\begin{aligned} r_u(0) &= \frac{\gamma \lambda_e \lambda_v \kappa p}{N}, \\ r_{uv}(0) &= -\frac{\gamma \lambda_e \lambda_v \kappa}{N}. \end{aligned} \quad (7)$$

We are now ready to make the following observations:

- To guarantee that the identified parameters satisfy the application requirement with high probability α , we require $\frac{N}{\lambda_e \kappa} - \gamma > 0$. Hence the highest possible achievable accuracy for a given noise variance, experimental length and probability is

$$\gamma < \frac{N}{\lambda_e \kappa}.$$

Or conversely, to achieve a certain accuracy the experimental length need to satisfy

$$N > \gamma \lambda_e \kappa.$$

- As expected the required input power during the identification increases with higher probability (larger κ), higher noise variance (larger λ_e) and tighter application requirements (larger γ), while increasing the experimental length, N , reduces the required power.
- Not as obvious is that the required input power is proportional to the disturbance variance. One might think that a higher variance of v_t could help us to identify the parameter corresponding to the disturbance filter, H , and hence require less accuracy in the identification of G , consequently requiring less power in u_t . While it is true that a higher power of the disturbance makes the identification of H more accurate, the application cost is also proportional to λ_v . Hence higher accuracy is needed to be able to satisfy the application requirements.
- The required input power is proportional to p . The constant p captures the effect of the underlying system,

$$p = \frac{h_0^2 - 2h_0 b_0 + 1}{1 - b_0^2} = \frac{(h_0 - b_0)^2}{1 - b_0^2} + 1 \geq 1.$$

Hence a lower bound on the required input power is given by

$$r_u(0) \geq \frac{\gamma \lambda_e \lambda_v \kappa}{N}.$$

- If the two systems are equal, i.e., $b_0 = h_0$, then $p = 1$. Thus the least amount of power in the input is needed if the two systems are equal. In system identification of structured systems it is recognized that it could be hard to identify subsystems which are equal, see [10],

[11]. Here we see that in this particular scenario this is not the case. If h_0 and b_0 are very different from each other p will be large and the required power is large. Furthermore we can see that if b_0 is close to ± 1 then high power is required. This is expected since we invert G_0 in our feed-forward filter and in this case we are close to the stability margin and a high accuracy is needed.

- If it is known beforehand that the two subsystems are equal this should of course be taken into account. In this case the feed-forward filter becomes $F_f = -1$ and is independent of the identified parameters. Hence it is not necessary to perform any system identification.
- The optimal input signal should always be negatively correlated with the disturbance and be

$$r_{uv}(0) = -\frac{\gamma\lambda_e\lambda_v\kappa}{N},$$

independent of the underlying system.

B. Non-minimum phase system

Now we look at the case when G_0 is a non-minimum phase system, i.e., $|b| > 1$. The optimal feed-forward controller on the form (5) will now be

$$F_f(q, \theta) = \frac{(hb^2 + b - h)q + hb}{b^2(bq + 1)}.$$

Even if we know the parameters of the true system the variance will not be zero in this case. The smallest error we can make is given by

$$V_{min} = \frac{(b_0^2 - 1)(b_0 - h_0)^2}{b_0^4}.$$

The application cost is now given by

$$V_{app}(\theta) = E \left\{ ((H + G_1 F_f) v_t)^2 \right\} - V_{min} \leq \frac{1}{\gamma}.$$

Note that $1/\gamma$ does no longer correspond to the highest acceptable variance but is now a bound on how much higher the variance can be using the identified parameters compared to if we knew the true system. Redoing the same calculations as in the minimum phase case, the optimal input signal is if

$$\frac{N}{\lambda_e\kappa} - \gamma \frac{b_0^4 - b_0^2 + 1}{b_0^4} \geq 0,$$

given by

$$\begin{aligned} r_u(0) &= \frac{\gamma\lambda_e\lambda_v\kappa}{N} p_2, \\ r_{uv}(0) &= -\frac{\gamma\lambda_e\lambda_v\kappa}{N} \frac{h_0(b_0^4 - 3b_0^2 + 4) + 2b_0^3 - 3b_0}{b_0^5}, \end{aligned}$$

where

$$\begin{aligned} p_2 &= \frac{h_0^2(b_0^6 + 16b_0^2 - 6b_0^4 - 10)}{b_0^6(b_0^2 - 1)} \\ &+ \frac{h_0(4b_0^5 - 18b_0^3 + 12b_0) + 4b_0^4 - 3b_0^2}{b_0^6(b_0^2 - 1)}. \end{aligned} \quad (8)$$

Many of the observations for the minimum phase setting also hold in this case, the differences are

- To achieve a certain accuracy the experimental length must now satisfy

$$N > \gamma\lambda_e\kappa \frac{b_0^4 - b_0^2 + 1}{b_0^4}.$$

- The required input power is proportional to p_2 defined in (8). We see that, for a fixed h_0 , if $|b_0| \rightarrow \infty$ then

$$p_2 \rightarrow 0 \quad \Rightarrow \quad r_u(0) \rightarrow 0.$$

Thus a larger b_0 requires less input power to achieve a certain accuracy.

- If $b_0 = h_0$ then

$$\begin{aligned} r_u(0) &= \frac{\gamma\lambda_e\lambda_v\kappa}{N} \frac{b_0^4 - b_0^2 + 1}{b_0^4}, \\ r_{uv}(0) &= -\frac{\gamma\lambda_e\lambda_v\kappa}{N} \frac{b_0^4 - b_0^2 + 1}{b_0^4}. \end{aligned}$$

- Again if $|b_1| \rightarrow 1$ then $r_u(0) \rightarrow \infty$. This is at the stability margin of our feed-forward controller and it is required that we identify G correctly as minimum phase or as non-minimum phase. Thus high accuracy and consequently high input power is needed.

C. Signal generation

We will now show how we can realize a input signal from the optimal correlations. Here we will only show how to do this for the minimum-phase case as the calculations for the non-minimum phase case is analogous. From the optimal correlations (7) we can generate an input signal by using

$$u_t = K v_t + r_t,$$

where r_t is a white zero mean Gaussian process independent of v_t with variance λ_r . The constant K can in this case be seen as a feed-forward filter that should be used during the identification. Since

$$\begin{aligned} r_u(0) &= K^2 \lambda_v + \lambda_r, \\ r_{uv}(0) &= K \lambda_v, \end{aligned}$$

we obtain

$$\begin{aligned} K &= \frac{r_{uv}(0)}{\lambda_v} = -\frac{\gamma\lambda_e\kappa}{N}, \\ \lambda_r &= r_u(0) - K^2 \lambda_v = \lambda_v \left(\frac{\gamma\lambda_e\kappa p}{N} - \frac{\gamma^2 \lambda_e^2 \kappa^2}{N^2} \right) \geq 0, \end{aligned}$$

where the last inequality is due to that $N > \gamma\lambda_e\kappa$ to have a feasible optimization problem.

D. White noise input

Let us compare the results with what we get if we use a white input signal during the experiment, uncorrelated with the measured disturbance v_t , i.e., $r_{uv}(0) = 0$. Now the optimization problem becomes

$$\begin{aligned} &\underset{r_u, r_{uv}}{\text{minimize}} && r_u(0) \\ &\text{subject to} && \frac{N}{\lambda_e\kappa} \begin{bmatrix} r_u(0) & 0 \\ 0 & \lambda_v \end{bmatrix} \succeq \gamma\lambda_v \begin{bmatrix} p & -1 \\ -1 & 1 \end{bmatrix} \end{aligned} \quad (9)$$

with optimal input

$$r_u(0) = \frac{\gamma\lambda_e\lambda_v\kappa p}{N} + \frac{\gamma^2\lambda_e^2\kappa^2\lambda_v p}{N(N - \gamma\lambda_e\kappa)}.$$

Comparing to the case with correlated input the required power is

$$\frac{\gamma^2\lambda_e^2\kappa^2\lambda_v p}{N(N - \gamma\lambda_e\kappa)}$$

larger, see (7). The difference decays as $1/N^2$ so for large N one could instead use an input signal uncorrelated with v_t .

IV. THE GENERAL CASE

We will now study more general cases, when the sub-systems are not necessary FIR-filters. Mainly we will focus the discussion on how to guarantee that the spectrum we optimize over, Φ_{χ_0} actually defines a spectrum, and how to formulate this as an LMI. Two common ways to do this are to use a finite dimensional parametrization or a partial correlation parametrization of the spectrum.

In the finite dimensional parametrization the spectrum is written as an infinite series and the optimization parameters are given by the truncated vector of coefficients in the series. Using the Kalman-Yakubovich-Popov lemma the condition that the parameters represent a spectrum can be written as an LMI, see [12] for details.

Often the optimization problem (4) can be written in terms of only a finite number of parameters. The idea of the partial correlation approach is to find conditions on these parameters that guarantee the existence of an infinite extension such that the complete sequence defines a spectrum. This can also be expressed as an LMI see [12].

The two above approaches are however not directly applicable to our problem. The problem is that we do not have control over the second input v_t . For example, the partial correlation approach only guarantees that there exist an extension, not that the extension exactly corresponds to the given properties of v_t .

The two approaches have however been extended to closed loop optimal input design, see for example [13] for the finite dimensional parametrization and [14] for the partial correlation approach. We will show that our problem is just a special case of this. In closed loop input design the objective is to find the controller K and the spectrum of the reference signal r_t , with $u_t = -Ky_t + r_t$, to achieve some properties on the identified models. See Figure 2a. This is the same as designing the spectrum of u , Φ_u and the cross spectrum between u_t and s_t , Φ_{us} . This is exactly what we want to do in the feed-forward case, we want to find conditions such that Φ_u and Φ_{us} corresponds to realizable experiment. We see that the feed-forward case in Figure 2b corresponds to the closed loop case when $G_0 = 0$ and $H_0 = -M$. Hence we can use the existing results from closed loop optimal input design to guarantee that the spectra in the optimization can be realized by a reference signal Φ_r and a controller K with $u_t = -Ky_t + r_t = Kv_t + r_t$. Here we will show how the

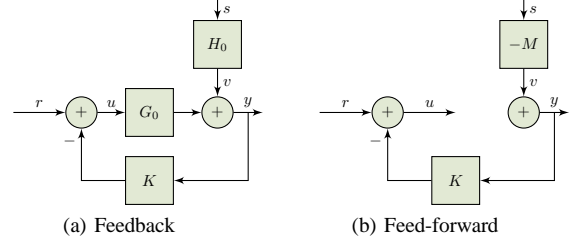


Fig. 2: The relation between the input u_t in the closed loop case and the feed-forward case.

partial correlation approach from closed loop optimal input design [14] can be used in this setting.

First we define the generalized moments of the spectrum Φ_{χ_0} as

$$m_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1}{|d(e^{j\omega})|^2} \Phi_{\chi_0} e^{jk\omega} d\omega, \quad (10)$$

where $d(z) = \sum_{l=0}^m d_l z^l$ is such that the coefficients are real and obey $d_m \neq 0$ and has all its roots outside the closed unit disc. Note that m_k is real and that $m_{-k} = m_k^T$.

Many common cost functions in input design, for example minimum input variance, can be written as linear functions of the generalized moments, see [14]. If the model structures of $G(q, \theta)$ and $H(q, \theta)$ are rational then by choosing $d(z)$ as the least common denominator of $\frac{\partial G(z, \theta)^*}{\partial \theta}$, $\frac{\partial H(z, \theta)^*}{\partial \theta}$ and $M^*(z)$ evaluated at $\theta = \theta_0$ we see that P^{-1} is affine in the generalized moments m_0, m_1, \dots, m_n c.f., (1), see [12] for details.

The input design problem (4) can hence be written as a SDP in a finite number of generalized moments (10). Now we want to classify the set of all sequences m_0, \dots, m_n that corresponds to a realizable experiment design, i.e., that corresponds to a K and Φ_r that are realizable. In [14] a semi-definite description of this set is given. We will shortly summarize the result here.

The sequence (m_0, m_1, \dots, m_n) of 2×2 matrices satisfying $m_{-k} = m_k^T$ for $k = 0, \dots, n$ defines a valid experiment if the following conditions hold¹

- 1) $m_{k,22} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\lambda_s}{|d(e^{j\omega})|^2} e^{jk\omega} d\omega$ for $k = -n, \dots, n$.
- 2) $\sum_{l=0}^m d_l m_{k-l,21} = 0$ for $k = 1, \dots, n$.
- 3) The matrix

$$T_n = \begin{bmatrix} m_0 & m_1^T & \cdots & m_n^T \\ m_1 & m_0^T & & m_{n-1}^T \\ \vdots & & \ddots & \vdots \\ m_n & m_{n-1} & \cdots & m_0 \end{bmatrix}$$

is positive semi-definite.

A proof and more details can be found in [14].

Remark 2: The above formulation is not only useful for feed-forward cases. It could as well be applied, with minor modifications, to more general MIMO system identification problems where some of the inputs are known but not

¹ $m_{k,pq}$ denotes the (p, q) element of the matrix m_k .

controllable. This is a common case in many industrial applications.

A. Generating the Input Signal

Solving the optimization problem gives the optimal m_0, m_1, \dots, m_n . The problem is then to find the feed-forward controller K and the spectrum Φ_r from these matrices.

Define the matrix $T_n^r(a) = \text{diag}(a, 0, a, 0, \dots, a, 0)$ and let a_{max} be the largest possible value of a such that $T_n - T_n^r(a)$ is positive semi-definite. Furthermore let \mathbf{v} be a non-zero vector such that $(T_n - T_n^r(a_{max}))\mathbf{v} = 0$ with $\mathbf{v} = (p_n, q_n, p_{n-1}, q_{n-1}, \dots, p_0, q_0)^T$. Then one possible realization of the reference spectrum and the feed-forward controller is

$$\Phi_r(\omega) = |d(e^{j\omega})|^2 a_{max},$$

$$K = -\frac{q(z)}{p(z)M(z)},$$

where $p(z) = \sum_{l=0}^n p_l z^{-l}$ and $q(z) = \sum_{l=0}^n q_l z^{-l}$. The spectrum of r can thus be realizable by filtering white noise with variance a_{max} through the filter $d^*(z)$. Other realizations are possible, see [14] for details and a proof.

V. EXAMPLE - HOUSE HEATING

In this section we will show how the framework for input design for feed-forward control can be applied in a simulation example. We will consider the problem of controlling the temperature in a room using electrical radiators. The control signal u_t , is the temperature of the radiator and y_t is the temperature in the room that we would like to control. The temperature in the room is also influenced by the outside temperature, v_t , due to heat transfer in the walls. To simplify the problem we consider v_t as white noise with variance λ_v . A thermometer is fitted outdoors so we assume that we can measure the temperature outside perfectly. The model of the system we will use here is a simplified version of the one given in [15]. The model is given by

$$y_{t+1} = -by_t + k_1 u_t + k_2 v_t$$

where b, k_1 and k_2 are some constants that depend on the heat transfer coefficients between radiator and air, the coefficient through the walls and the sampling time. This can be rewritten as

$$y_t = \underbrace{\frac{k_1}{1+bq^{-1}}}_{G} u_t + \underbrace{\frac{k_2}{1+bq^{-1}}}_{H} v_t + e_t$$

where we have added e_t as zero mean white Gaussian measurement noise with variance λ_e .

The objective is to identify the parameters $\theta = (k_1, k_2, b)^T$ and design a feed-forward filter such that we keep the variance in the indoor temperature due to changing outdoor temperature less than $1/\gamma$.

The feed-forward filter based on the identified parameters will be

$$F_f(q, \theta) = -\frac{k_2}{1+bq^{-1}} \left(\frac{k_1}{1+bq^{-1}} \right)^{-1} = -\frac{k_2}{k_1}.$$

Since both G and H have the same dynamics we see that the accuracy of the the parameter b is not important. The application cost is hence

$$V_{app}(\theta) = E \{ ((H_0 + G_0 F_f(\theta))v_t)^2 \} \leq \frac{1}{\gamma}$$

which can be calculated analytically using for example residue calculus. From this we can then derive $V_{app}''(\theta_0)$.

We will now use the framework outlined in Section IV to formulate an optimal experiment for this case. First we look at the gradient \mathcal{F}_0 defined in (2)

$$\mathcal{F}_0 = \begin{bmatrix} \frac{\partial G(q, \theta)}{\partial \theta} & \frac{\partial H(q, \theta)}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \frac{1}{1+bq^{-1}} & 0 \\ 0 & \frac{1}{1+bq^{-1}} \\ -\frac{k_1 q^{-1}}{(1+bq^{-1})^2} & -\frac{k_2 q^{-1}}{(1+bq^{-1})^2} \end{bmatrix}$$

and

$$P^{-1} = \frac{1}{2\pi\lambda_e} \int_{-\pi}^{\pi} \mathcal{F}_0 \Phi_{\chi_0} \mathcal{F}_0^* d\omega.$$

By defining $d(e^{j\omega}) = (1 + be^{j\omega})^2$ and m_k as in (10) we can write P^{-1} as linear combinations of m_k , $k = -2, \dots, 2$.

The input signal energy can be written in terms of m_k as

$$\begin{aligned} \frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_u(\omega) d\omega &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{|d(e^{j\omega})|^2}{|d(e^{j\omega})|^2} \Phi_u(\omega) d\omega \\ &= \sum_{k=-2}^2 \delta_k m_{k,11}, \end{aligned} \quad (11)$$

where δ_k are the coefficients of $|d(e^{j\omega})|^2$, i.e., $|d(e^{j\omega})|^2 = \sum_{k=-2}^2 \delta_k e^{jk\omega}$. In the same way we can express the output variance during the experiment as

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \Phi_y(\omega) d\omega = \sum_{k=-1}^1 \beta_k \begin{bmatrix} k_1 & k_2 \end{bmatrix} m_k \begin{bmatrix} k_1 \\ k_2 \end{bmatrix}, \quad (12)$$

where β_k are the coefficients of $|1 + be^{-j\omega}|^2$.

We can now formulate the optimal input design problem where we want to find either a minimum variance input signal or a input signal that gives the lowest output variance while satisfying the application constraints with high probability. The optimization becomes

$$\text{minimize}_{m_k, k=-2, \dots, 2} \quad (11) \text{ or } (12)$$

$$\text{subject to} \quad \frac{N}{\kappa} P^{-1} \succeq \gamma V_{app}''(\theta_0)$$

$$m_{k,22} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\lambda_v}{|d(e^{j\omega})|^2} e^{jk\omega}, \quad \forall k$$

$$\sum_{l=0}^m d_l m_{k-l,21} = 0, \quad k = 1, \dots, 2$$

$$\begin{bmatrix} m_0 & m_1^T & m_2^T \\ m_1 & m_0 & m_1^T \\ m_2 & m_1 & m_0 \end{bmatrix} \succeq 0.$$

The above problems are difficult, if not impossible, to solve analytically. As the problems are SDPs they can be solved efficiently using numerical methods. The following system parameters will be used; $b = -0.5$, $\lambda_v = 3$, $\lambda_e = 1$, $k_1 = 1$

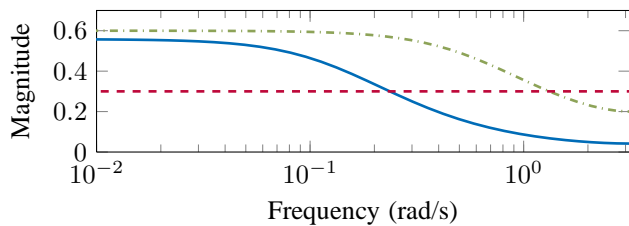


Fig. 3: The magnitude of the optimal feed-forward filters to be used during the identification for the minimum input variance case (—) and the minimum variance output (---). For reference H_0 is also shown as (-.-).

and $k_2 = 0.3$. The experimental conditions are $N = 1000$, $\gamma = 100$ and $\kappa = 5.99$ corresponding to at least 95% probability that the identified model satisfies the application requirement.

The two problems are solved with CVX, a package for specifying and solving convex programs [16], giving the optimal m_k , $k = -2, \dots, 2$. Using the algorithm outlined in Section IV-A we generate a reference spectrum Φ_r and the feed-forward controller to be used during the identification, K . In both cases the optimal T_n is singular and consequently $a_{max} = 0$ and $\Phi_r = 0$. Therefore it is enough to use $u_t = Kv_t$ during the identification. The magnitude of the optimal feed-forward controllers K for both problems are shown in Figure 3 where also H_0 is shown for reference. We see that when we minimize the input variance the optimal K is a lowpass filter, while for the output variance case the optimal K is a constant $K = -0.3$. The case when $K = -0.3$ corresponds to the optimal feed-forward controller $F_f = -k_2/k_1 = -0.3$ that gives zero output variance in y due to the measurable disturbance v_t . But using this feed-forward filter during the identification gives that $y_t = e_t$, and hence we cannot say anything about the parameters k_1 , k_2 and b from this information. However, since we know that the subsystems have the same dynamics, this information is enough to identify the ratio k_2/k_1 which is all information we need to design our feed-forward controller!

We verify the results in a Monte-Carlo simulation. In each round system identification data is generated using the optimal identification feed-forward filter, K , and the parameters k_1 , k_2 and b are identified. The identified parameters are then used to design the feed-forward controller. Finally we calculate the output variance due to the measurable disturbance v_t when the feed-forward filter $F_f(\hat{\theta}_N)$ is used in the application. In about 98% of the simulations the output variance was less than $1/\gamma$, thus the results seem to be valid.

VI. CONCLUSIONS

In this paper we have considered optimal input design when the identified model will be used in a feed-forward control application. First, the first order FIR-filter case was considered and some fundamental properties were observed. Secondly a framework for optimal input design for feed-forward systems was presented and the relation to closed

loop input design was discussed. The framework was then successfully applied to a numerical example.

Interesting extensions of this work would be to see what happens if, on top of the feed-forward, feedback is added as this is the most common case in practice. It would also be interesting to see if some of the results for the low order FIR-case can be extended to more general structures.

We noted that the framework presented here can be extended to more general MIMO system identification problems where we only can control a few of the inputs. It would be interesting to see if this can be used in identification of complex interconnected systems, see [17]. For example if one node locally want to identify the dynamics of the network, how should it excite the system when it only can measure, and not affect, the inputs from its neighbors?

REFERENCES

- [1] M. Gevers, "Identification for control: From the early achievements to the revival of experiment design," in *Decision and Control, 2005 and 2005 European Control Conference. CDC-ECC '05. 44th IEEE Conference on*, Dec., pp. 12–12.
- [2] M. Gevers and L. Ljung, "Optimal experiment designs with respect to the intended model application," *Automatica*, vol. 22, no. 5, pp. 543 – 554, 1986. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/0005109886900646>
- [3] X. Bombois, G. Scorletti, M. Gevers, P. V. den Hof, and R. Hildebrand, "Least costly identification experiment for control," *Automatica*, vol. 42, no. 10, pp. 1651 – 1662, 2006. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0005109806002238>
- [4] H. Hjalmarsson, "System identification of complex and structured systems," *European Journal of Control*, vol. 15/3-4, pp. 275–310, 2009.
- [5] B. Wahlberg, H. Hjalmarsson, and M. Annergren, "On optimal input design in system identification for control," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, Dec., pp. 5548–5553.
- [6] C. Larsson, M. Annergren, and H. Hjalmarsson, "On optimal input design in system identification for model predictive control," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, Dec., pp. 805–810.
- [7] L. Ljung, *System Identification: Theory for the User, 2nd Edition*. Upper Saddle River, New Jersey: Prentice Hall, 1999.
- [8] K. Åström and B. Wittenmark, *Computer controlled systems: theory and design*, ser. Prentice-Hall information and system sciences series. Prentice-Hall, 1984.
- [9] B. Wahlberg, H. Hjalmarsson, and P. Stoica, "On the performance of optimal input signals for frequency response estimation," *IEEE Transactions on Automatic Control*, vol. 57, no. 5, pp. 766–771, 2012.
- [10] B. Wahlberg, H. Hjalmarsson, and J. Mårtensson, "Variance results for identification of cascade systems," *Automatica*, vol. 45, no. 6, pp. 1443 – 1448, 2009.
- [11] P. Hägg, B. Wahlberg, and H. Sandberg, "On identification of parallel cascade serial systems," in *18th IFAC World Congress*, vol. 18, 2011, pp. 9978–9983.
- [12] H. Jansson, "Experiment design with applications in identification for control," Ph.D. dissertation, Stockholm, 2004.
- [13] H. Hjalmarsson and H. Jansson, "Closed loop experiment design for linear time invariant dynamical systems via lmis," *Automatica*, vol. 44, no. 3, pp. 623 – 636, 2008. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0005109807003524>
- [14] R. Hildebrand, M. Gevers, and G. Solari, "Closed-loop optimal experiment design: The partial correlation approach," in *Decision and Control (CDC), 2010 49th IEEE Conference on*, Dec., pp. 2855–2862.
- [15] T. Glad and L. Ljung, *Control Theory*, ser. Control Engineering. Taylor & Francis, 2000.
- [16] I. CVX Research, "CVX: Matlab software for disciplined convex programming, version 2.0 beta," <http://cvxr.com/cvx>, Sep. 2012.
- [17] A. Dankers, P. Van den Hof, P. Heuberger, and X. Bombois, "Dynamic network identification using the direct prediction-error method," in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, Dec., pp. 901–906.